

TITLE OF THE INVENTION: Method and Apparatus for Data Replication using SCSI over TCP/IP

5

BACKGROUND OF THE INVENTION

Related Applications

This application claims priority from U.S. Provisional Application No. 60/206,607, filed May 23, 2000, incorporated herein by reference.

Field of Invention

10

The invention relates to replication of data in a computer system. More particularly, the present invention relates to utilizing the SCSI transport layer over TCP/IP, via an IP multicast mechanism, in order to replicate data transferred to a SCSI device.

Description of the Related Art

15

The Small Computer Systems Interface ("SCSI") is a popular family of protocols for communicating with I/O devices, in particular storage devices. More to the point, SCSI is the basic protocol for I/O operations between computers and storage subsystems. It is, in essence, a point-to-point protocol.

20

Another popular protocol is the Transport Control Protocol/Interface Program ("TCP/IP"). TCP/IP is the basic set of communication protocols for general data exchange between computers connected on a communication network. This is a common protocol used to communicate via the Internet.

Currently there is a convergence between the two protocols, that is, SCSI and TCP/IP. In particular, computers that communicate with their subsystems via SCSI are now tending to be interconnected via the Internet and utilizing TCP/IP to communicate with each other.

25

In view of this convergence, a standard has been proposed for implementing SCSI over TCP. The currently proposed standard "-draft-ietf-ips-iSCSI-06.txt" is available at

<http://www.ietf.org/internet-drafts/draft-ietf-ips-iscsi-06.txt> (herein expressly incorporated by reference) (the “Standard”). The Standard aims to be fully compliant with the requirements laid out in the SCSI Architecture Model-2 Document (also herein expressly incorporated by reference). This Standard provides for transmitting SCSI commands between SCSI devices, over
5 TCP/IP connections. Conventionally, SCSI devices that communicate with each other must be connected via the same SCSI bus. The proposed Standard permits SCSI devices which are not connected via the same SCSI bus to communicate with each other via the Internet, utilizing TCP/IP protocol. According to the Standard, SCSI devices that are separated from one another even by great distances may communicate with each other. The Standard seeks to describe an
10 implementation of SCSI that operates on top of TCP as the transport protocol.

There are numerous higher level storage functions that, in the context of a SCSI protocol, are implemented as a set of point-to-point I/O operations. One of the most well known of these higher level storage functions is Data Mirroring. In Data Mirroring on a conventional SCSI system, data replication is implemented by issuing separate SCSI I/O operations to each
15 device that is intended hold an image of the replicated data.

The following briefly defines some SCSI conceptual terms. The end point of a typical SCSI command is a “logical unit” (LUN). Logical units include, for example, hard drives, tape drives, CD and DVD drives, printers and processors. A collection of logical units is referred to as a “target” and is directly addressable on the network. In a client-server model, the target
20 corresponds to the server. An “initiator” creates and sends SCSI commands to the target. In the client-server model, the initiator corresponds to the client.

A typical SCSI command results in a command phase, data phase and a response phase. In the data phase, information travels either from the initiator to the target (for example, a WRITE command), or from the target to the initiator (for example, a READ command). In the
25 response phase, the target returns the final status of the operation, including any errors. A response signals the end of a typical SCSI command.

In a conventional storage subsystem, data mirroring is accomplished by providing additional logic, generally on the I/O operation level. Typically, for each WRITE operation for

data which is mirrored, this additional logic issues a corresponding WRITE operation to every device that maintains the mirrored image of the data. Thus, if a conventional system provides for two-way mirroring, two copies of each WRITE command are sent over the SCSI bus; moreover, two copies of the data associated with the WRITE command are sent over the SCSI bus,

5 resulting in multiple transfers of the same data. Consider, for example, that WRITE commands from an initiator such as a personal computer are to be mirrored to three targets such as three data storage devices, e.g., ZIP drives. Every time the initiator generates a WRITE command which is to be mirrored, the WRITE command and associated data must be sent to all three of the targets. Moreover, the I/O subsystem requires additional complicated logic in order to handle the data replication, error handling and data routing within the SCSI network. As can be seen from
10 foregoing, data mirroring typically dramatically increases the use of the SCSI bus. This, in turn, results in a significantly slower throughput to other devices utilizing the SCSI bus and increases the probability of errors.

Figure 1 is an illustration of a typical remote data mirroring system, shown in U.S. Patent
15 No. 6,173,377 B1, to Yanai et al. This system automatically provides and maintains a copy or mirror of data stored at a location geographically remote from the main or primary data storage device. The data processing system 110 in Figure 1 can be configured for remote mirroring from a user interface of the service processor 134 in the primary data storage system 114. Remote mirroring software 113 can be provided; host application programs can interface with a remote
20 mirroring facility of the data storage systems 114, 146 via the remote mirroring software 113. According to Yanai process and system, central processing unit 152 could be located with the remote secondary data storage system 146, linked to the primary and secondary data storage systems 114, 146 via redundant signal paths. Communication links 140, 141 from link adapters 136, 137 are provided. To provide remote mirroring operating modes for specific applications,
25 the remote mirroring facility defines an operating mode for each logical volume of data in the storage device as in the data storage systems 114, 146. A specific set of methods are provided in order to remotely mirror data according to either a synchronous or a semi-synchronous mode. When a host writes data to a remotely mirrored volume, a series of steps occurs. Data is written to the cache of the data storage system with the primary volume; an entry is queued up to
30 transmit data to the secondary volume; the secondary volume's data storage system

acknowledges receipt of the data; and a write end signal is returned to the host that initiated the WRITE request.

BRIEF SUMMARY OF THE INVENTION

5 It is the intent of this invention to provide a method, system and/or means for mirroring data transfer commands in a network, the network including a SCSI initiator and SCSI targets, the initiator and targets all being connected to a TCP/IP neetwork. Included is transmitting, from the initiator, a registration command via the TCP/IP network, to each target of the targets, the registration command including a multicast address and a virtual logical unit number. Also
10 included is establishing a mapping, responsive to the registration command, at each of the SCSI devices, indicating an association between the SCSI device and the virtual logical unit number and the multicast address. Further included is transmitting, from the initiator, a SCSI data transfer command, the SCSI data transfer command indicating the multicast group address. Also included is receiving, at each of the members of the multicast group, the data transfer command, and transferring data in response thereto. Further included is executing, in each of the SCSI
15 members of the multicast group, the specified data transfer command.

According to one embodiment of the invention, the invention includes transmitting, from the initiator, to each target of the targets, a mirroring termination command. According to an embodiment, the mirroring termination command causes a disassociation from the virtual logical
20 unit number; optionally, the mirroring termination command causes a disassociation from the multicast group.

According to a preferred embodiment of the invention, the data transfer command is a WRITE command. The WRITE command specifies the virtual logical unit number. A sequence may included multipleWRITE commands.

25 According to a preferred embodiment of the invention, the registration command is transmitted to a unicast address corresponding to each of the targets. Preferably, the registration command has a standard SCSI text command format.

According to a highly preferred embodiment, each of the targets responds to the registration command by setting a communication NIC to listen on the multicast address; and by mapping the virtual logical unit number to the SCSI device. Further, according to the highly preferred embodiment, upon establishing the mapping, each of the targets returns a message to the initiator.

According to a preferred embodiment, the data transfer command is transmitted via a multicast. Further according to a preferred embodiment, wherein each of the SCSI members of the multicast group returns a status of the data transfer command, and the initiator handles any error status of the data transfer command.

Moreover, according to another embodiment, each of the targets returns, to the initiator, a status indicating completion of the mirroring termination command.

The present invention proposes to take advantage of the IP multicasting mechanism, to simplify the logic to a great extent, to eliminate unnecessary data replication and to take advantage of general multicasting logic already implemented in most IP routers. These and other objects, features and advantages of the present invention are readily apparent from the following drawings and detailed description of the preferred embodiments.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

The above-mentioned and other advantages and features of the present invention will become more readily apparent from the following detailed description and the accompanying drawings, in which:

Figure 1 is a block diagram of a prior art system illustrating remotely mirrored primary and secondary data storage systems and links.

Figure 2 is a blocked diagram illustrating a SCSI initiator connected via the Internet (a TCP/IP backbone and appropriate routers) to targets.

Figure 3 is a diagram illustrating the registration command protocol.

Figure 4 is a diagram illustrating data transfer phase protocol.

Figure 5 is a diagram illustrating one version of data transfer error handling protocol.

Figure 6 is a diagram illustrating another version of data transfer error handling protocol.

Figure 7 is a diagram illustrating a third version of data transfer error handling protocol.

5 Figure 8 is a diagram illustrating the protocol for mirroring terminating phase protocol.

DETAILED DESCRIPTION OF THE INVENTION

The following detailed description of the preferred embodiment includes many specific details. The inclusion of such details is for the purpose of illustration only and should not be understood to limit the invention. Throughout this discussion, similar elements are referred to by similar numbers in the various figures for ease of reference.

Figure 2 is an exemplary embodiment of a SCSI initiator 201 that is connected via a TCP/IP backbone 203 including routers to three targets, target 0 205A, target 1 205B, and target 2 205C. According to highly preferred embodiments, the SCSI initiator 201 and the targets 205A, B and C are connected to the Internet as specified in the Standard.

Although for the purposes of illustration it assumed that the TCP/IP backbone including routers is part of the global Internet, it should be noted that the invention can be applied to any system utilizing SCSI buses which are interconnected via a TCP/IP protocol. Thus, the invention could easily be used on an Intranet, for example.

Typically, a target is connected via a SCSI bus to its own SCSI devices or logical units. In the example shown in Figure 2, target 0 205A includes two SCSI devices, namely logical unit 0 207 and logical unit 1 209; target 1 205B includes three SCSI devices, logical unit 0 211, logical unit 1 213 and logical unit 2 215; also included is target 2 205C including one SCSI device, namely logical unit 0 217. In this example, it is intended that logical unit 0 207 of target 0 205A, logical unit 1 213 of target 1 205B and logical unit 0 217 of target 2 205C are mirrored.

This example configuration will be used in the following discussion of the method according to the invention. Reference is made to Figure 2, in connection with describing the mirroring of all WRITE commands generated by the initiator 201 on three SCSI devices logical unit 0 207 on target 0 205A, logical unit 1 213 on target 1 205B and logical unit 0 217 on target 2 205C (the mirrored devices). More specifically, all WRITE commands from the initiator 201 will be executed by the mirrored devices.

The Standard accommodates the expansion of commands to include special commands which are not defined in the Standard. Nevertheless, in order to accommodate the special commands, the initiator and the SCSI devices need to be modified somewhat in order to recognize the special commands. The present invention is advantageously implemented in accordance with the Standard's provisions that accommodate using such special commands.

Accordingly, the SCSI devices utilized at the targets should be modified somewhat from a standard SCSI device. As a first point, the SCSI devices should be capable of connecting to a TCP/IP network, preferably as described in the Standard. Further, in order to use SCSI devices according to the invention, the devices should be modified to be capable of recognizing the special commands relating to data mirroring, and of executing subsequently received commands appropriately.

The present invention does not necessarily contemplate any alterations of the non-WRITE operations, or of any other operation defined in the Standard.

The invention envisions four stages, in order to implement the mirroring of WRITE operations to multiple targets:

Multicast group establishment phase.

Data transfer phase.

Response reception phase.

Mirroring termination phase.

Each of these phases will be discussed in greater detail below. It is possible to have multiple data transfer and response reception phases following the multicast group establishment phase, and prior to the mirroring termination phase.

Multicast Group Establishment Phase

Reference is made again to the example in Figure 2. The multicast group establishment phase involves the following steps, not necessarily in this order. First, each of the designated targets which is intended to be a part of the multicast group is associated with a multicast address. Next, each of the SCSI devices that are intended to be part of the multicast group are associated with a particular virtual logical unit number. In this manner, a command which is sent to the particular virtual logical unit number may be interpreted to be mirrored to each of the logical unit numbers which constitute a part of the multicast group.

For example, assume that targets 0, 1 and 2 of Figure 2 are intended to be associated with a multicast group. According to the first step, targets 0, 1 and 2 are associated with the multicast address. Second, the device with logical unit number 0 207 at target 0 205A, the device with logical unit 1 213 at target 1 205B and the device with logical unit number 0 217 at target 2 205C are associated with the same virtual logical unit number.

It will be appreciated that SCSI devices should be modified in order to be capable of recognizing a command to associate themselves with a virtual logical unit number, to execute subsequently received commands specifying that virtual logical unit number, and to recognize a command to disassociate themselves from the virtual logical unit number.

In accordance with the foregoing, part of the multicast establishment phase includes informing the SCSI devices that are intended to participate in the mirroring of a virtual logical unit number associated with the group. Subsequently, WRITE commands that are intended to be mirrored will specify the virtual logical unit number, and need not specify the individual actual logical unit numbers of the mirrored SCSI devices.

A portion of the multicast establishment phase has the purpose of informing all SCSI devices that are intended to participate in the mirroring of a virtual logical unit number

associated with a group. All subsequent WRITE commands that are to be mirrored will specify the virtual logical unit number.

Once the multicast group has been established, any message sent by the initiator specifying the multicast address or the virtual logical unit number will be distributed to all targets that are part of that group. Also, all WRITE operations (discussed in detail below) that specify the virtual logical unit number of the multicast group will be executed by all SCSI devices associated with that virtual logical unit number.

The multicast establishment phase is intended to be executed at the beginning of a sequence of commands which are intended to be mirrored. Advantageously, the multicast group establishment phase is executed only once. However, the multicast establishment phase could be re-executed, as might be desirable, for example in order to change the definition of the group. For example, a re-transmission of the multicast establishment phase could result in adding or deleting targets from the group.

Reference is made to Figure 3, showing the preferred protocol for a multicast group establishment phase. Multicast group establishment is accomplished according to the following steps:

1. The initiator transmits a SCSI registration command, via the TCP/IP connection, to each target that will participate in the mirroring. As illustrated, the initiator transmits a registration command 301 to target 1; transmits a registration command 303 to target 2; and continues until it transmits the final registration command 305 to the final target n. Each of these registration commands is sent to the Unicast address of each one of the targets.

According to highly preferred embodiments, the registration command is implemented using the standard SCSI text command as defined in the "Standard" section 2.8, reproduced below.

The standard SCSI text command includes a text field. In the text field, the registration command specifies the following information, as illustrated above: the multicast group IP

address, the SCSI logical unit of the SCSI device to which the command is sent, and the SCSI virtual logical number to be used by future WRITE commands that are to be mirrored.

The following illustrates an example of the text field of preferred registration command for a target with three logical units (luns) to participate in a mirror session:

5 Command=register
 multicast_group_ip = X.X.X.X\0
 lun_map=(lun1#,lun2#,lun2#),vlun#)\0

Note that all three SCSI devices on the same target have to be mapped to the same virtual device. Consider the example shown in Figure 2. Assume that the multicast group ip address to be used is 128.12.15.3. The registration command sent to target 1 will specify, as illustrated below, the virtual logical unit number, and the logical unit number one, thereby informing target 1's device with logical unit 1 that it should associate itself with the specified virtual logical unit number. The following text portion in the registration command uses the above-illustrated format for the just-discussed example:

10 Command=register\0
 multicast_group_ip = 128.12.15.3\0
 lun_map=(1,5)\0
 mapping lun 1 to virtual lun 5
15

20 As was indicated above, SCSI devices participating in the mirroring according to this invention should be modified to recognize and operate under the registration command.

It is possible for multiple SCSI devices at a single target to become part of the group. In that instance, the registration command would specify more than one logical unit number, so that all specified SCSI devices at that target associate themselves with the virtual logical unit number.

25 Alternatively, multiple registration commands, one for each SCSI device, can be directed to the target with the multiple SCSI devices which are part of the group.

2. Responsive to the registration command, the targets will respond by registering to the multicast group using the group IP address. As shown in Figure 3, target 1 registers to the multicast group 307, target 2 registers to the multicast group 309, and so on until target n which also registers to the multicast group 311. A target may register to the multicast group by: 1) setting the communication NIC to listen on the specified multicast address; and 2) setting the target's internal (implementations specific) data structures to map the virtual lun number with the real devices.

3. In response to the registration command, SCSI devices within a target will establish a mapping between the SCSI virtual logical unit number, as specified by the registration command, and the device's real SCSI logical unit number. As shown in Figure 3, in response to the registration command passed through the target, the SCSI device at target 1 establishes the specified mapping 313, the targeted SCSI device of target 2 establishes the specified mapping 315 and so on until the final SCSI device at the final target n, which also establishes the specified virtual logical unit number to logical unit number mapping 317.

4. Upon completion of the actions performed to become responsive as part of the multicast group, each target preferably sends a message back to the initiator that it has completed the registration process. According to highly preferred embodiments, each of the targets will send the response after the target has registered to the multicast group and the specified SCSI devices at the target have established the appropriate mapping. Preferably, the message is returned as a SCSI text response frame back to the initiator as specified in section 2.9 of the Standard, and reproduced below.

As illustrated in Figure 3, each target sends the registration confirmation message to the Unicast address of the initiator. Thus, target 1 returns a registration confirmation 319 to the initiator, target 2 returns a confirmation 321 to the initiator, and so on until the final target n returns a registration confirmation 323 to the initiator.

The multicast group establishment phase is complete when the initiator has received a confirmation from all targets in the multicast group confirming that all targets and all SCSI devices on those targets that have been specified are registered with the group.

According to one alternative embodiment, the same group of targets may participate in multiple mirroring sessions however on different devices. In this alternative, the same group of targets are included in multiple multicast groups, however, typically a different initiator will be responsible for each different multicast group. Referring back to the example of Figure 2, the first mirroring session can include one group of devices including the device at target 0 with logical unit number 0, the device at target 1 with logical unit number 1, and the device at target 2 with logical unit number 0. The second mirroring session may consist different devices. The first group of devices may mirror WRITE commands from one initiator in the first mirroring session, and the second group of devices would mirror WRITE commands from another initiator as the second mirroring session. In this example, the multicast address for the first and second groups can be the same since the same targets are included in both groups. Nevertheless, the virtual logical unit number for the first group should be different from the virtual logical unit number for the second group.

According to another alternative, the same group of targets are participating in multiple mirroring sessions. However, according to this alternative, both groups of devices mirror WRITE commands from the same initiator.

In yet another alternative, multiple initiators can participate in the same mirroring session. In this case however the initiator registering the session needs to inform the other initiators through TCP/IP or any other communication means of the multicast address and the virtual logical unit for the session.

In conventional systems, when targets register to a group or associate themselves with a multicast address, it is the target that must initiate the registration process. Thus, in a traditional system, if an initiator was transmitting some audio visual data such as a movie, and a target wanted to receive the data, it was up to the target to register or associate itself with the multicast address. However, according to the present invention, the initiator is responsible for establishing the group. That is, by sending the registration commands to pre-determined targets, the initiator controls which targets join the group.

Data Transfer Phase

The data transfer phase is executed every time the initiator sends a WRITE command to the multicast address. This assumes that the multicast group has been established. During the data transfer phase, the data which is being written is transferred from the initiator to each of the targets which are part of the multicast group.

According to the data transfer phase, a WRITE command is transmitted to the multicast group, resulting in a WRITE I/O to each of the logical units that constitute a part of the multicast group. The data transfer phase is preferably done according to the following steps:

1. The initiator sends a SCSI WRITE 401 command to the multicast group IP address, that is the group established by the multicast group establishment phase. The command is multicast to all targets participating in the mirroring. Advantageously, this is done the same way that any standard SCSI command would be multicast, preferably in accordance with the Standard. In the illustrated preferred protocol for the data transfer phase, there is a single WRITE command 401 being sent from the initiator to the entire multicast group.

A typical write command can be found in the Standard in section 2.7, reproduced below.

2. Each target member of the multicast group then returns a ready status indication to the initiator, advantageously via a standard ready to transfer ("RTT"), preferably sent as a SCSI response in accordance with the Standard. The RTT should be sent from every target which was configured as part of the multicast group. Thus, as illustrated in Figure 4, target 1 returns a RTT response 403, target 2 returns a RTT response 405 and so on through the final target n which also returns a RTT response 407. The initiator waits until it receives a RTT response from every target, prior to sending the actual data transmission itself.

3. Having received a RTT from each of the targets, the initiator then sends the data portion of the command 409 to the multicast group IP address and the corresponding virtual logical unit number. As is illustrated, the data is thereby transferred to multiple SCSI devices using a single data command.

In accordance with highly preferred embodiments, the initiator can obtain the status of the WRITE command from each member of the multicast group during the response phase, discussed below.

Errors in transmitting all of the data maybe handled in any of several ways.

5 Conventionally, there is no standard way of ensuring that each member of a group in a conventional system receives every data block sent by an initiator to a multicast address. Loss of data may be unimportant, such as in the case of transmission of audio/visual data, for example, a movie or video conference. In such situations, it is not fatal if a target occasionally does not receive one of the transmitted blocks. In those situations, missing a block could merely cause the
10 target to lose the signal for some fraction of a second, and moreover receiving a re-transmitted block could have undesirable results. Nevertheless, when truly mirroring data WRITE commands, for example in providing redundant data storage, it is essential that every target receives every data block transmitted by the initiator, in order to maintain the integrity of the transmitted data.

15 The standard SCSI WRITE command specifies the amount of data that is going to be written. As a result, in the preferred embodiment, a SCSI device is aware of precisely how much data it should receive as part of any WRITE command. If any SCSI device receives less data than what is expected from a WRITE command, it can report to the initiator that all data was not received.

20 Reference is made to Figure 5. Here, it is illustrated that a WRITE command 401 was transmitted from the initiator to the multicast group. In this particular example, each of the targets experience a data failure. Thus, target 1 returns a data failure report 501, target 2 returns a data failure report 503, and so on through the final target n which also returns a data failure report 505. In this situation, it is appropriate for the initiator to re-transmit the entire WRITE
25 command 507 to the entire multicast group.

More typically, in a data transfer error, it is just one of the targets or one of the SCSI devices at one of the targets that has experienced data failure. Reference is made to Figure 6. Here, the WRITE command 401 was transmitted from the initiator to the multicast group. This

WRITE command dropped data to one of the targets. Target 1, which experienced the data failure, then returns a data failure report 601 to the initiator. The initiator then re-transmits the WRITE command 603 as a unicast to the target that indicated it had dropped the data. This can be repeated for additional targets that have experienced dropped data.

5 Reference is made to Figure 7. As an alternative to handling errors in the data transfer phase, and avoiding re-transmitting the WRITE command, it may be desirable to re-transmit only those data blocks that were not received. According to this alternative, the initiator labels or numbers each data block of a transmission. SCSI devices receiving the data then note the numbers or labels of the data blocks as they are received. If less than the total amount of data or
10 data blocks were received by the SCSI device during a WRITE command, the SCSI device missing data can determine which block or blocks were not received and request a re-transmission of only those missing data blocks 701. The initiator then re-transmits just the missing data blocks 703 via unicast to the SCSI device that requests the re-transmission. According to this alternative, SCSI devices would be modified to be capable of monitoring data
15 block numbers and labels.

The foregoing are given by way of example for handling of data transmission errors. Nevertheless, other solutions for ensuring that every target in the group receives all of the data blocks transmitted by the initiator will be apparent to those skilled in the art.

Response Reception Phase

20 It is advantageous to have a response reception phase executed following each transmission by the initiator of a WRITE command to the multicast address. During the response reception phase, each of the targets reports status information associated with the WRITE command back to the initiator. The initiator may then determine if the WRITE command was successful at the multicast group level.

25 Reference is made back to Figure 4. Following the data transfer phase, the response reception phase is performed. It is done in the following steps:

1. Conventionally, SCSI devices always report the status of a WRITE command to the initiator. According to the preferred embodiment, each of the SCSI devices indicates the status of the multicast WRITE command in the same way it would report the status of a conventional WRITE command. During the response reception phase, each SCSI device associated with the virtual logical unit number specified by the WRITE command will send their status information to the initiator. Thus, as illustrated, following the data portion of the WRITE command 409, the SCSI device at target 1 transmits a status response 411, the SCSI device at target 2 transmits a status response 413, and so on through the final SCSI device at target n which also transmits a status response 415 to the initiator. Preferably, these are conventional status responses and are sent via unicast to the initiator's TCP/IP address. The responses need not be multicast.

The initiator itself should wait until it receives a status response 411, 413, 415 from all the targets in the multicast group. Provision can be made to time out in the event that a status response does not arrive from a SCSI device at a particular target.

2. If all responses received from all the targets indicate a successful WRITE completion, then the initiator may return a "successful operation" indication to the application 417. On the other hand, if one of the responses indicates a failure or is not received, then the error should be handled. One possible action is to return a status to the application from the initiator, wherein the status indicates a bad WRITE and uniquely identifies the failed target, thereby allowing the application itself to take any recovery action it may choose. Alternatively, the initiator itself may retry the WRITE operation to the failed target using its unicast TCP address.

The response reception phase could be omitted in certain embodiments, particularly in situations where the status of the data transfer is inconsequential. However, in most data mirroring situations, it is anticipated that application program will want to be aware of the successful completion (or lack thereof) of each write operation.

Mirroring Termination Phase

The mirroring termination phase is used to disband the multicast group. More specifically, it disassociates the targets from the multicast address and disassociates the SCSI devices from the virtual logical unit number. It should be necessary to execute this phase only once following the establishment and completion of the multicast group WRITE (or series of WRITES). After execution of the mirroring termination phase, WRITE commands from the initiator will not be mirrored to members of the multicast group.

Reference is made to Figure 8. The termination phase is performed according to the following steps:

1. The initiator sends a command indicating that the mirroring phase is terminated 801. Preferably, this is done via a standard text SCSI command, and indicates the mirroring termination phase on the particular multicast group address. This command is also preferably based on the SCSI command. According to the preferred embodiment, the text field of the mirroring termination command will specify both the multicast group address and the virtual logical unit number which are to be terminated.

An example format for the preferred embodiment of the text portion for the mirroring termination command follows:

```
command=terminate\0
```

```
virtual_lun=vlun#\0
```

Each target receiving the mirroring termination command will unregister itself from the multicast group, further removing the association between its logical unit numbers and its virtual logical unit number for this particular session. As is illustrated in Figure 8, the first target unregisters itself 803, the second target unregisters itself 805, and so on through the final target n which also unregisters itself 807. Unregistration does the following two things: 1) removes the mapping between the devices and the virtual device from the target's internal data structures; and

2) if this is the last mirroring session for the target, instructs the communication NIC to stop listening on the multicast IP address.

3. Upon completion of the unregistration tasks, each of the targets will then send a text response frame, preferably as specified in section 2.9 of the Standard (reproduced below), indicating completion of the mirroring termination phase to the initiator. Thus, upon completion of unregistration, target 1 returns an unregistration complete text response frame 809 to the initiator, target 2 returns a similar unregistration complete text response frame to the initiator 811, and so on through the final target n which also returns an unregistration complete response frame 813 to the initiator.

This invention has been described in connection with specific examples, which are intended to illustrate the invention. The invention, nevertheless, encompasses many variations on the described examples. For example, although the invention has been described in connection with three example targets, it is equally applicable to more or fewer targets. Further, although the illustrated example includes one, two or three SCSI devices on each target, the invention is equally applicable where there are more than three logical units on each target. Also, the illustrated example presumes that the initiator communicates with the targets through a TCP/IP backbone including routers; it is technically possible to implement the invention without utilizing a TCP/IP backbone and/or a router. Further, certain examples assume that the data transfer command that is mirrored is a WRITE command; the multicast may be equally applicable to other data transfer commands which are not technically WRITE commands.

Moreover, although the preferred embodiment of the invention utilizes the SCSI/TCP formats established in the Standard, it is possible to implement the commands so that they are not in conformance with the Standard, so long as the SCSI devices are adapted to recognize such commands. Also, one of skill in the art will recognize that there are other appropriate methods for error recovery in addition to those outlined herein; the invention is intended to encompass such error recovery methods. The protocols have been illustrated with targets 1, 2 ... n, although the invention is equally applicable to any number of targets, including just one target. Specific methods for registering to the multicast group, and for establishing mapping have been provided

by way of example; equally appropriate methods will be apparent and are encompassed by the invention. Also, the exchange of transmissions according to the protocols has been illustrated in an orderly fashion; the sequence of transmissions may occur in a different order and still be within the scope of the invention.

- 5 Sections 2.7, 2.8 and 2.9 of the current version of the Standard are reproduced below for reference:

2.7 SCSI Data-out & SCSI Data-in

The typical data transfer specifies the length of the data payload, the Target Transfer Tag provided by the receiver for this data transfer, and a buffer offset. The typical SCSI Data PDU for WRITE (from initiator to target) has the following format:

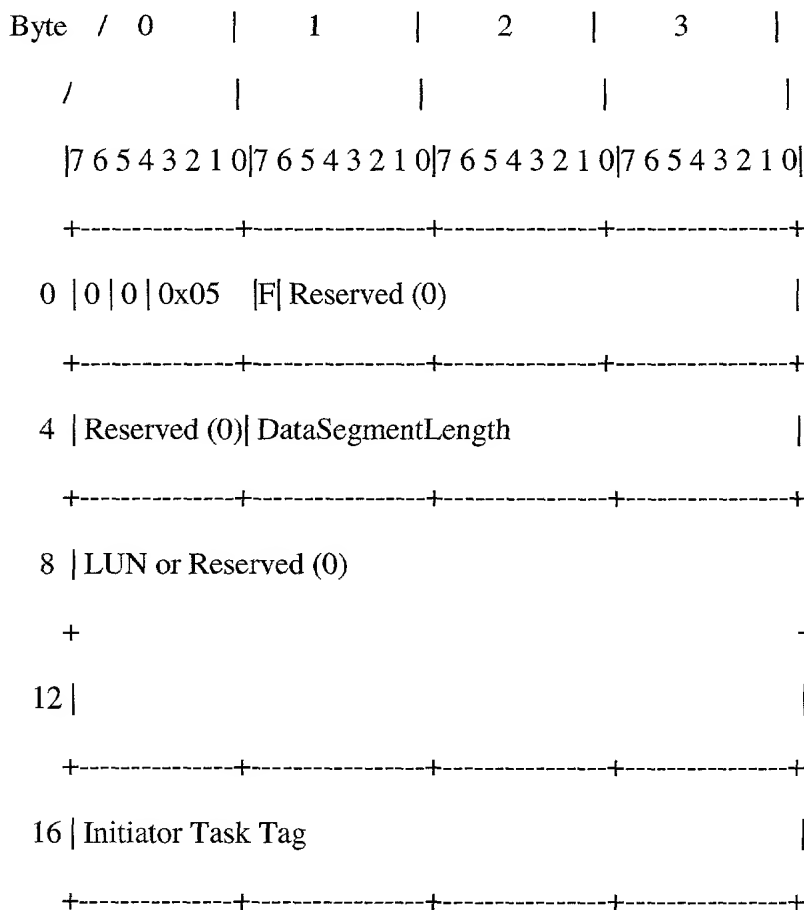
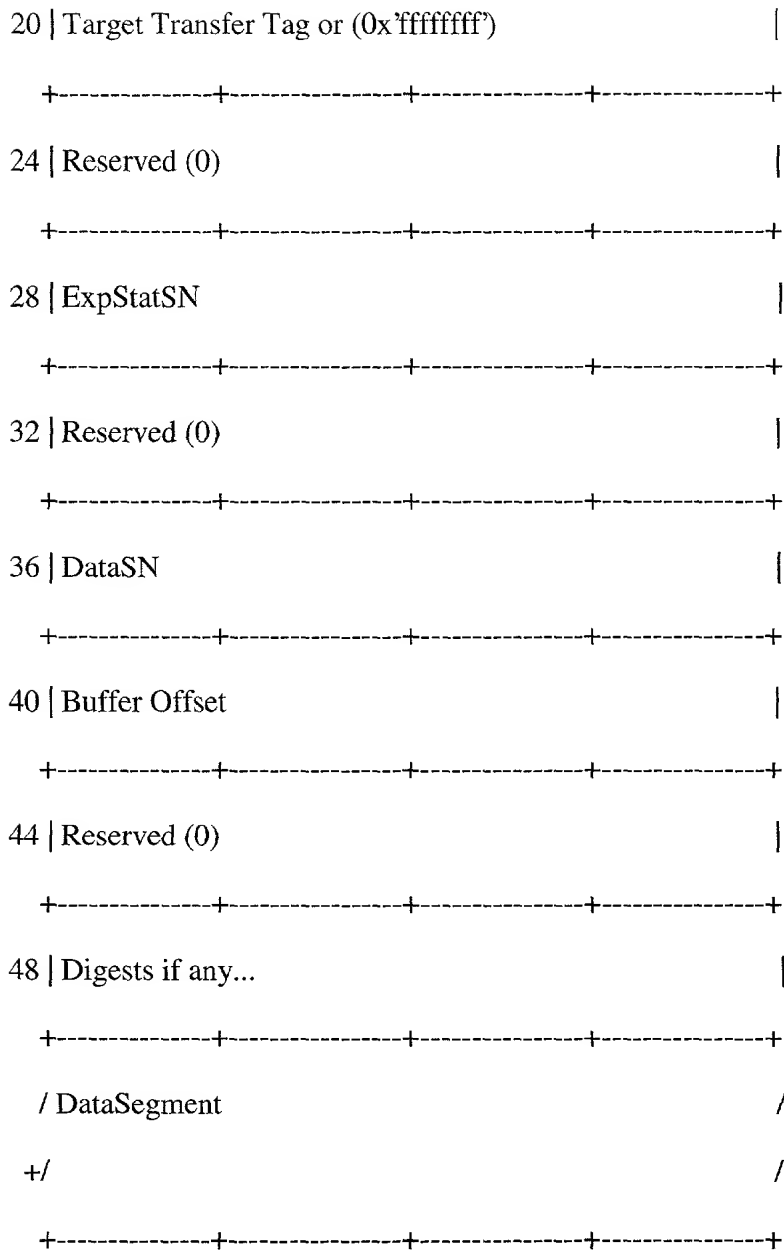
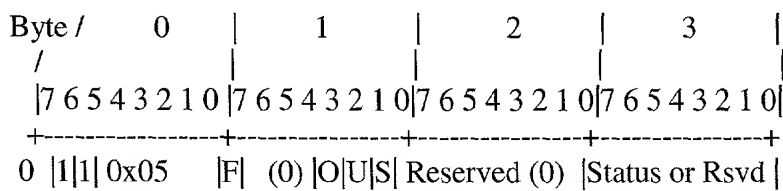
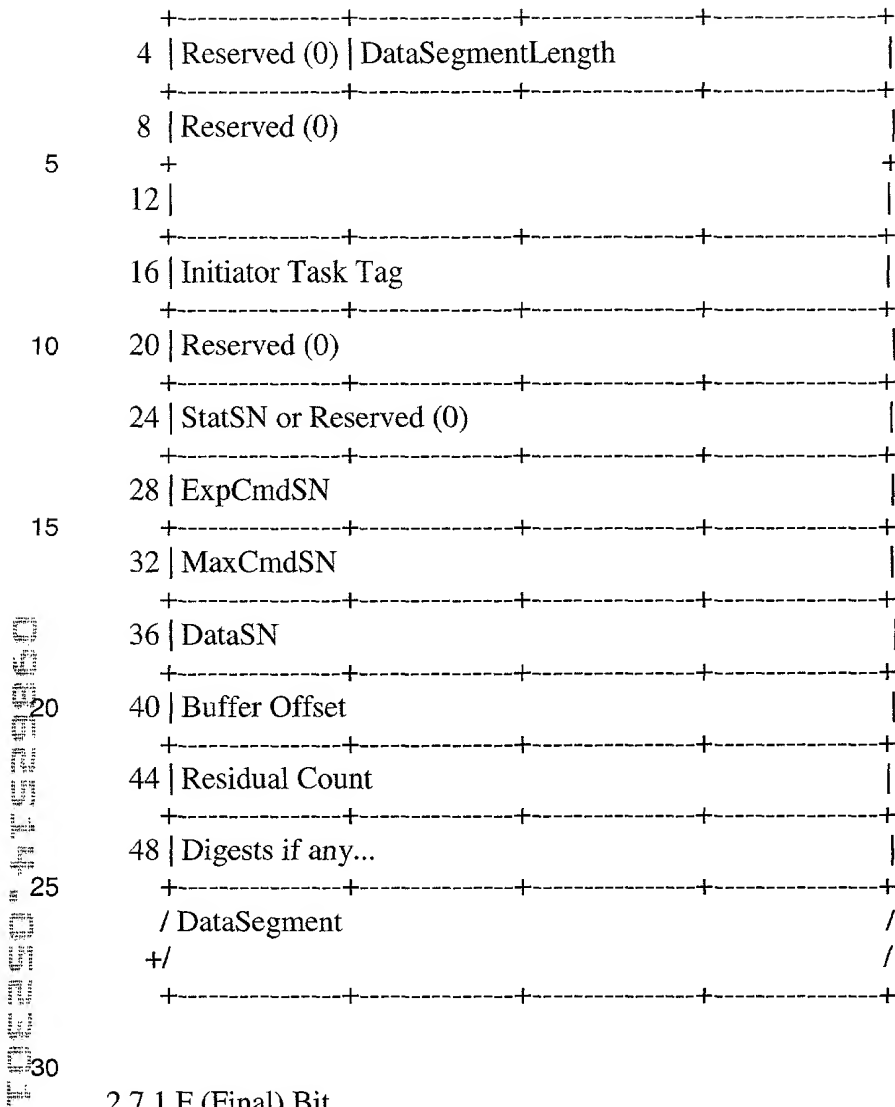


Figure 10



The typical SCSI Data packet for READ (from target to initiator) has the following format:





2.7.1 F (Final) Bit

For outgoing data, this bit is 1 for the last PDU of unsolicited data or the last PDU of a sequence answering a R2T. For incoming data, this bit is 1 for the last input data PDU associated with the command (even if it includes the status).

2.7.2 Target Transfer Tag

On outgoing data, the Target Transfer Tag is provided to the target if the transfer is honoring a R2T. In this case, the Target Transfer Tag field is a replica of the Target Transfer Tag provided with the R2T.

The Target Transfer Tag values are not specified by this protocol except that the all-bits-one value (0x'ffffff') is reserved and means that the Target Transfer Tag is not supplied. If the Target Transfer Tag is provided then the LUN field MUST hold a valid value and be consistent with whatever was specified with the command, otherwise the LUN field is reserved.

2.7.3 StatSN

This field MUST be set only if the S bit is set to 1.

2.7.4 DataSN

For input (read) data PDUs, the DataSN is the data PDU number (starting with 0) within the data transfer for the command identified by the Initiator Task Tag. For output (write) data PDUs, the DataSN is the data PDU number (starting with 0) within the current output sequence. The current output sequence is identified by the Initiator Task Tag (for unsolicited data) or is a data sequence generated for one R2T (for data solicited through R2T).

Any input or output data sequence MUST contain less than $2^{32}-1$ numbered PDUs.

2.7.5 Buffer Offset

The Buffer Offset field contains the offset of this PDU payload data against the complete data transfer. The sum of the buffer offset and length should not exceed the expected transfer length for the command.

Input data ordering is governed by a disconnect-reconnect mode page bit (EMDP). If this bit is 0 the target MUST deliver packets in increasing buffer offset order.

Output data within a burst (initial or any data PDU sequence that fulfils a R2T) MUST be delivered in increasing buffer offset order.

2.7.6 DataSegmentLength

This is the data payload length of a SCSI Data-In or SCSI Data-Out PDU; sending of 0 length data segments should be avoided.

2.7.7 Flags

The last SCSI Data packet sent from a target to an initiator for a particular SCSI command that completed successfully may also optionally contain the Command Status for the data transfer. In this case, Sense Data cannot be sent together with the Command Status. If the command is completed with an error, then the response and sense data MUST be sent in a SCSI Response PDU (i.e., MUST NOT be sent in a SCSI Data packet). For Bi-directional commands the status MUST be sent in a SCSI Response PDU.

b0 S (status)- set to indicate that the Command Status field contains status. If this bit is set to 1 the F bit MUST also be set to 1

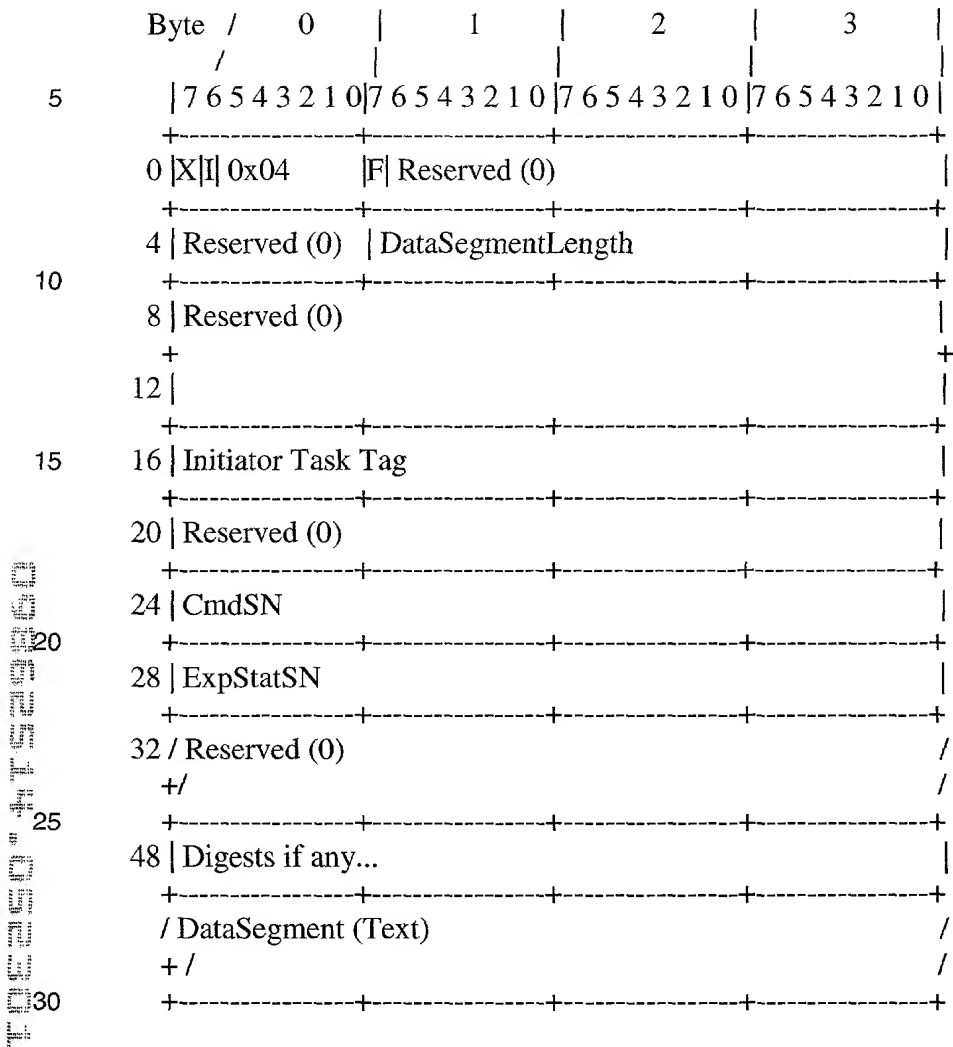
b1-2 as in an SCSI Response

b3-6 not used (should be set to 0)

The fields StatSN, Command Status, Residual Count have meaningful content only if the S bit is set to 1.

2.8 Text Command

The Text Command is provided to allow the exchange of information and for future extensions. It permits the initiator to inform a target of its capabilities or to request some special operations.



2.8.1 F (Final) Bit

When set to 1 it indicates that this is the last or only text command in a sequence of commands; otherwise it indicates that more commands will follow.

2.8.2 Initiator Task Tag

The initiator assigned identifier for this Text Command.

If the command is sent as part of a sequence of commands (e.g., the Login Phase or a sequence of Text commands) the Initiator Task Tag **MUST** be the same for all the commands within the sequence (similar to linked SCSI commands).

5 2.8.3 Text

The initiator sends the target a set of key=value or key=list pairs encoded in UTF-8 Unicode. The key and value are separated by a '=' (0x3D) delimiter. Many key=value pairs can be included in the Text block by separating them with null (0x00) delimiters. A list is a set of
10 values separated by comma (0x2C). Large binary items can be encoded using their hexadecimal representation (e.g., 8190 is 0x1FFE) or decimal representation. The maximum length of an individual value (not its string representation) is 255 bytes.

The data length of a text command or response **SHOULD** be less than 4096 bytes. No
15 key **SHOULD** contain more than 255 characters.

Character strings are represented as plain text. Numeric and binary values are represented using either decimal numbers or the hexadecimal 0x'ffff' notation. The result is adjusted to the specific key.
20

The target responds by sending its response back to the initiator. The response text format is similar to the request text format.

Some basic key=value pairs are described in Appendix A and D. All of these keys, except
25 for the X- extension format, **MUST** be supported by iSCSI initiators and targets.

Manufacturers may introduce new keys by prefixing them with X- followed by their (reversed) domain name, for example the company owning the domain acme.com can issue:

30 X-com.acme.bar.foo.do_something=0000000000000003

Any other key not understood by the target may be ignored without affecting basic function. If the Text Response does not contain a key that was requested, the initiator must assume that the key was not understood by the target or, whenever appropriate, that the response was "none".

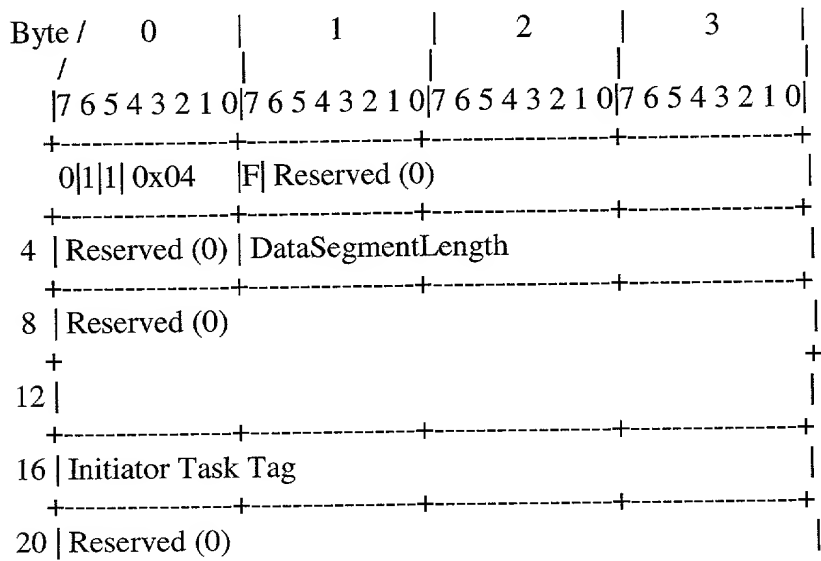
Text operations are usually meant for parameter setting/negotiations but can be used also to perform some active operations.

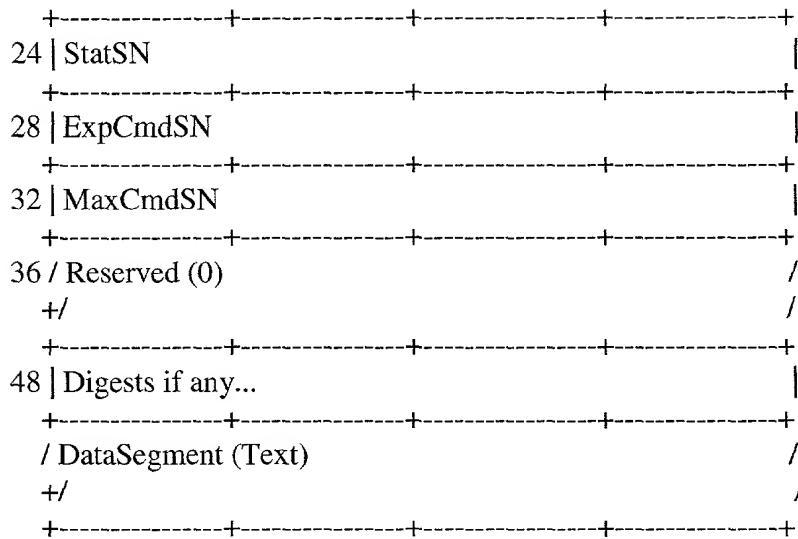
It is recommended that Text operations that will take a long time should be placed in their own Text command.

A session may have only one outstanding text command or text command sequence at any given time.

2.9 Text Response

The Text Response PDU contains the target's responses to the initiator's Text Command. The format of the Text field matches that of the Text Command.





2.9.1 F (Final) Bit

When set to 1 in response to a text command with the Final bit set to 1 the F bit indicates that the target has finished it's operation. Otherwise if set to 0 in response to a text command with the Final Bit set to 1 it indicates that the target has more work to do (invites a follow-on text command). A text response with the F bit set to 1 in response to a text command with the F bit set to 0 is a protocol error.

2.9.2 Initiator Task Tag

The Initiator Task Tag matches the tag used in the initial Text Command or the Login Initiator Task Tag.

2.9.3 Text Response Data

The Text Response Data Segment contains responses in the same key=value format as the Text Command and with the same length and coding constraints. Appendix C lists some basic Text Commands and their Responses. If the Text Response does not contain a key that was

requested, the initiator must assume that the key was not understood by the target or that the answer is <key>=none.

- Text response key=value pairs MUST be delivered in the same order as the command
- 5 key=value pairs whenever applicable.

- While this invention has been described in conjunction with the specific embodiments outlined above, many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, the preferred embodiments of the invention is set forth above are
- 10 intended to be illustrative, and not limiting. Various changes may be made without parting from the spirit and the scope of the invention as defined in the following claims.

11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 265